**Identify online patient conversations**

An IT team collaborates with the Social Listening team for automating the process of gaining insights from social media conversations.

The Social Listening team has to manually validate heart failure related conversations fetched from the social listening tool which scans twitter, Facebook, forums, blogs etc.  Such conversations are posted by multiple stakeholders like patients, doctors, media houses, general public, etc. The team needs to identify the patient conversations, so as to dig deeper into them and identify the patient needs. The data science team wants to automate this process by building intelligent algorithms to predict patient conversations.

Build an Intelligent pipeline that can segregate patient conversations from the rest of the group given historically tagged patient data.

You are expected to build an algorithm where they can ingest the social data and get the patient tags - 1 if patient and 0 if not a patient.

Description of attributes in dataset is given below

|  |  |
| --- | --- |
| **Header** | **Header Description** |
| Source | Type of Social Media Post |
| Host | Domain of Social Media Post |
| Link | Complete URL of post |
| Date | Date of Post |
| Time | time Stamp of Post in Eastern Time |
| Time(GMT) | time Stamp of Post in GMT |
| Title | Title of the Post |
| TRANS\_CONV\_TEXT | Actual Text Conversation of the Post |